

### **REMARKS/ARGUMENTS**

This communication is responsive to the Office Action mailed December 15, 2005.

A terminal disclaimer over U.S. Pat. App. No. 10/058,684 is enclosed herewith.

Claim 51 has been amended to clarify the claim and over the 112 rejection.

Claim 33 has been amended to overcome the anticipation rejection.

### **Claims 1-4, 29, and 31-32**

Claims 1-4, 29, and 31-32 stand rejected under 35 USC 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (“Kawashima”) in view of Okayama et al. (US 6,556,767) (“Okayama”).

The Examiner notes that Kawashima fails to disclose an event being characterized in a manner free from using a model sequence of frames to determine the approximate frame that the ball is put into play. In particular, Kawashima explicitly teach the using of a model sequence of frames to determine the approximate frame that the ball is put into play. This is a technique that is based upon attempting to model the characteristics of a baseball game. In contrast, claim 1 claims a technique that does not attempt to model the characteristics of a baseball game when the ball is put into play and considered out of play, but rather relies on other non-model based techniques.

The Examiner suggests that Okayama discloses characterizing an event in a manner free from using a model sequence of frames. Referring to FIG. 4, Okayama discloses the comparison of non-compressed video streams to other stored non-compressed video frames. In this occasion, differences between pixels are calculated by pixel-by-pixel comparison between the two non-compressed motion picture data, and the sum of the differences between all pixels is obtained to conduct comparison, so that a judgment is made as to whether there is any scene change or not (step 132). When the sum of the differences is larger than a predetermined threshold, a decision can be made that there is some scene change. See, column 11, lines 52 to column 12, lines 30. Accordingly, Okayama disclose an estimation of pixel changes to determine a scene change

The combination of Kawashyima et al. and Okayama would result in a system that identifies every single scene change in the video. However, there would be no indication of which scene changes are the start time and the end time, as claimed. In this manner, the result

would be an extensive series of scene changes, many of which are before the start time, one of which may be the start time (as suggested by the Examiner), many of which are after the start time, one of which may be the end time (as suggested by the Examiner), and many after the end time.

However, there is no suggestion that Okayama would be applicable to effectively summarizing baseball because there is no mechanism free from using a model sequence of frames to identify which scene changes indicate the start time and the end time, there is no suggestion that Okayama would be applicable to identifying events in baseball, nor is there any suggestion to modify Kawashima with a non-model based technique.

Claim 1 patentable distinguishes over the cited references by claiming the identification of events of a video that includes baseball, and in particular the time when a ball is put into play and the ball is considered out of play wherein the start time is identified before the end time and then subsequently the end time is next identified without identifying any other potential end times before the end time.

Claims 2-4 depend from claim 1 and are, therefore, patentable for at least the same reasons asserted for claim 1.

Claim 29 patentably distinguishes over Kawashima in view of Okayama et al. by claiming the end of the at least one of the segment is identified subsequently to a start time of the segment without identifying any other potential end times before the end time.

Claims 30-32 depend from claim 29 and are, therefore, patentable for at least the same reasons asserted for claim 29.

### **Claims 5, 7 and 8**

The Examiner rejected claims 5, 7 and 8 as being rejected under 103(a) as being unpatentable over Kawashima et al. in view of Lin et al., U.S. Patent No. 6,724,933.

Kawashima et al. disclose a system that use model sequence of frames, as previously noted. The Examiner notes that Kawashima et al. fail to disclose not using a model sequence of frames, not using an action spotting technique, and not using a pixel-by-pixel comparison.

Lin et al. disclose a media analysis tool for the processing of video. However, there is no suggestion in Lin et al. of how to analyze the content to identify the claimed segments.

Claim 5 patentably distinguishes over the cited references by claiming, in part, identifying a plurality of segments of the video based upon a series of activities defined by the rules of baseball that could potentially result in a least one of (1) a score, (2) preventing a score, (3) advancing a team toward a score, and (4) preventing advancing a team toward a score, where the identifying a plurality of segments of the video is in a manner free from using a model sequence of frames, free from using an action spotting technique, and free from using a pixel-by-pixel comparison.

Claim 5 further patentably distinguishes over the cited references by claiming, wherein the identifying identifies the start of one of the segments and the end of the one of the segments free from identifying any other potential ends of the one of the segments between the start and the end.

Claims 6-8 depend from claim 5 and are, therefore, patentable for at least the same reasons asserted for claim 5

### **Claim 9**

The Examiner rejected claim 9 under 103(a) as being unpatentable over Kawashima et al. in view of Okayama et al. and further in view of Lin et al.

As previously noted Okayama et al. and Lin et al. both fail to suggest how to accurately identify the desired frames, nor which characteristics or relationships should be used to make such an identification. Rather the combination of Okayama et al. and Lin et al. would result in the identification of a myriad of frames.

Claim 9 patentably distinguishes over the cited references by claiming the that the time frame put into play is identified before the end of the play and then subsequently the end of the play is next identified without identifying any other potential end of the play before the end of the play.

Claims 10-12 depend from claim 9 and are, therefore, patentable for at least the same reasons asserted for claim 9.

### **Claims 13-26 and 51**

Claims 13-26, and 51 stand rejected under 35 USC 103(a) as being unpatentable over Kawashima et al. in view of "Multimedia Content Analysis", Wang et al. ("Wang") and further in view of Lin et al.

Kawashima discloses a system for indexing baseball telecast for content-based video retrieval. In particular, Kawashima discloses a first stage that detects domain specific scene in baseball video based-on image similarity, referred to as a basic scene. The first stage is determined based upon using a similarity measure between an input image and a model image, together with a threshold. In a second stage, the exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image. Accordingly, Kawashima discloses using a similarity measure between a pair of frames and a model sequence of frames to determine the approximate frame that the ball is put into play, again based upon a set of model sequence. In essence, Kawashima teaches a technique based upon interpreting the video containing baseball in order to identify portions of interest.

With respect to claim 13, the Examiner states that Kawashima identifies a plurality of video segments based upon color characteristics of the video. The applicant would note that, section 1.1 relates to indexing of a baseball game by the recognition of an at bat and the identification of the player, which merely state a desirable goal for Kawashima but is unrelated to how that goal is achieved, and is unrelated to color. Section 1.1 has nothing to do with color characteristics of the video. Section 1.2 relates to the summarization of the video using indexes which merely states the desirability of reducing the amount of database store and has nothing to do with color characteristics of the video. Section 2.1 relates to the segmentation process which is based, in part, on the use of model-based matching (step 2 of figure 1) and action spotting techniques (step 3 of figure 1). Section 2.1 further discloses determining the first frame of the next show based upon a similarity measurement using color histograms (step 4). Accordingly, the technique uses model-based matching and action spotting techniques to detect a basic scene, with the first frame of the next shot determined based upon using a color histogram. However, the applicant respectfully suggests that the color histogram do not include any particular spatial information regarding the location of particular colors within the video.

The Examiner comments that section 2.1.3 suggests that color characteristics are included wherein the threshold is based upon pixels that are colored to form an image produced on the screen. The applicant would note that 2.1.3 teaches the use of continuous dynamic programming with a minimal warp function and pitching/batting model sequence. The resulting characteristic vector is based upon motion in the attention area I from pixels whose intensity changes in successive frames greater than a threshold. This characteristic vector fails to take into account continuous spatial regions of generally homogeneous color, but rather seeks to identify motion above a threshold of colored pixels.

It is possible to develop highly sophisticated models of a typical baseball video to identify potentially relevant portions of the video, as taught by Kawashima. However, such highly sophisticated models are difficult to create and are not normally robust. Further, the likelihood that a majority of the highly relevant portions of the baseball video will be included in such a video summarization is low because of the selectivity of the model. Thus the resulting video summarization of the baseball game may simply be unsatisfactory to the average viewer.

After consideration of the difficulty of developing highly sophisticated models of a baseball video to analyze the content of the baseball video, as the sole basis upon which to create a baseball summarization, the present inventors determined that this technique is ultimately flawed as the models will likely never be sufficiently robust to detect all the desirable content. Moreover, the number of different types of model sequences of potentially desirable content is difficult to quantify. In contrast to attempting to detect particular model sequences, the present inventors determined that the desirable segments of the baseball game are preferably selected based upon a "play." A "play" may be defined as an sequence of events defined by the rules of baseball. In particular, the sequence of events of a "play" may be defined as the time generally at which the ball is put into play (e.g., a time based upon when the ball is put into play) and the time generally at which when the ball is considered out of play (e.g., a time based upon when the ball is considered out of play). Normally the "play" would include a related series of activities that could potentially result in a score (or a related series of activities that could prevent a score) and/or otherwise advancing the team toward scoring (or prevent advancing the team toward scoring).

The present inventors then considered how to detect a “play” from a baseball video in a robust, efficient, and computationally effective manner. After extensive analysis of a typical baseball game it was determined that a baseball game is usually captured by cameras positioned at fixed locations around the baseball field, with each camera typically capable of panning, tilting, and zooming. Each play in a baseball game normally starts with the pitcher releasing the ball, such as toward the catcher or toward one of the basemen. Further, a pitching scene, in which the pitcher is about to throw the ball, is usually captured from a camera location behind the pitcher. This camera angle is typically used because it is easier to observe the movements of all of the parties involved (the pitcher, the batter, the catcher, and the umpire) from this viewpoint.

It is based upon these observations regarding the characteristics of a baseball broadcast and realization that part of the problem is the inability of complex sophisticated models to be sufficiently robust and computationally efficient that a spatial region of generally homogenous green color and a spatial region of a generally homogeneous brown color should be used. Without any such realizations, there would be no motivation to modify Kawashima to include color characteristics.

Claim 13 patentably distinguishes over Kawashima in view of Wang by claiming the start of the plurality of segments is identified based upon detecting at least one spatial region of a generally homogenous green color in a generally lower region of the video and at least two spatial regions of a generally homogenous brown color in a pair of spaced apart horizontal regions where the generally homogenous green region is between the horizontal regions.

Claims 14-18 depend from claim 13 and are, therefore, patentable for at least the same reasons asserted for claim 13.

As previously described, it is based upon those observations regarding the characteristics of a baseball broadcast and realization that part of the problem is the inability of complex sophisticated models to be sufficiently robust and computationally efficient that a different approach is desirable. Based upon a different approach it was determined that three horizontally oriented regions of generally homogenous colors including at least one of generally brown and generally green is beneficial. Without any such realizations, there would be no motivation to modify Kawashima to include spatial color characteristics of the regions.

Claim 19 patentably distinguishes over Kawashima in view of Wang and further in view of Lin et al. by claiming identifying a plurality of segments of the video, wherein the start of said plurality of segments is identified based upon detecting at least three spatial regions of horizontally oriented regions of generally homogenous colors including at least one of generally homogenous brown and generally homogenous green, where each of the segments includes a plurality of frames of said video.

Claims 20-23 depend from claim 19 and are, therefore, patentable for at least the same reasons asserted for claim 19.

The applicant would respectfully suggest that section 2.1.3 of Kawashima uses count of pixels whose intensity change in successive frames are larger than a threshold. There is no use of color characteristics disclosed by Kawashima to identify the play itself. Moreover, the Examiner suggests that somehow a more robust scene extraction technique would result from using color clusters. However, there is no suggestion in ANY of the cited references that the use of scene clusters would provide a more robust result for baseball video. Rather, the Examiner is improperly using hindsight to construct the claimed invention.

As previously described, it is based upon those observations regarding the characteristics of a baseball broadcast and realization that part of the problem is the inability of complex sophisticated models to be sufficiently robust and computationally efficient that a separate model is desirable. Without any such realizations, there would be no motivation to modify Kawashima to include spatial color characteristics to determine the start of a segment in baseball.

Claim 24 patentably distinguishes over Kawashima in view of Wang by claiming identifying a plurality of segments of the video, identifying a plurality of segments of said video, wherein the start of said plurality of segments is identified based upon identifying at least three separate spatial regions of sufficient spatial coherence and sufficient horizontal extent in a generally lower region of the video, where each of the segments includes a plurality of frames of the video.

It is noted that Wang, page 33 lines 27-46, relates to color histograms, color layout, shape, etc. However, there is no suggestion in Kawashima that there would be any desirability of the claimed regions for determining the start of a segment of baseball.

Claims 25-26 depend from claim 24 and are, therefore, patentable for at least the same reasons asserted for claim 24.

The Applicants would respectfully suggest that section 2.1.3 of Kawashima uses count of pixels whose intensity change in successive frames that are larger than a threshold. There is no use of color disclosed by Kawashima to identify the play itself.

The Examiner notes that Kawashima fails to disclose the use of textures. The applicant would respectfully suggest that Kawashima uses an intensity comparison technique that does not analyze spatial regions of the image to determine the start of a segment. Accordingly, there would be no motivation to include any type of spatial textural regions in the system of Kawashima to identify the start of a segment.

Claim 51 patentably distinguishes over Kawashima in view of Wang by claiming identifying wherein the start of the plurality of segments is identified based upon detecting different separate spatial regions having generally homogenous different textures.

#### **Claims 27 and 28**

Claims 27 and 28 stand rejected under 35 USC 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (“Kawashima”) in view of “Multimedia Content Analysis”, Wang et al. (“Wang”), further in view of Okayama et al. (“Okayama”) and further in view of Dimitrova et al. (“Dimitrova”).

Kawashima detects the start of a frame based upon the modeling technique for a sequence of motion events (see Figure 1, steps 1 through 3). There is no suggestion in Kawashima to modify its technique to remove candidate frames as the start of a play by comparison of a single frame to another temporally relevant single frame. Rather, Kawashima relies on model-based matching and action spotting to determine the start of a play, which are both multi-frame based techniques.

Claim 27 patentably distinguishes over the art of record by claiming identifying a plurality of segments of the video including a play, wherein the identifying for at least one of the



segments includes detecting the start of the segment of a play based upon processing of a first single frame of the video independently of other frames and removing the single frame as the start of said segment including a play based upon processing of another single frame temporally relevant to the first single frame independently of other frames.

Claim 28 depends from claim 27 and is, therefore, patentable for at least the same reasons asserted for claim 27.

### **Claim 33**

Claim 33 stands rejected under 35 USC 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (“Kawashima”) in view of Wang et al. (“Wang”).

Kawashima discloses a system for indexing a baseball telecast for content-based video retrieval. In particular, Kawashima discloses a first stage that detects a basic scene. The exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image. A set of frames are modeled against a similarity measure, and a minimal warp function is used to compare the input video sequence with pitching/batting model sequences. Thereafter, the end of the segments is identified from a rough estimate of the camera viewing angle of the next scene cut just after the pitching. See Kawashima, section 2.1.4. Accordingly, after identifying the start of the segment using the model sequence of frames the next scene cut is used to identify the end of the segment, all of which is based upon a model sequence of frames.

By contrast, claim 33 claims in part, detecting more than one scene change during a baseball video segment, such as multiple scene changes during a play of baseball, and identifying as the end of the segment, or end of the play, a scene change other than the first scene change during the segment. That is, claim 33 patentably distinguishes over Kawashima, at least in part, by claiming identifying a plurality of segments of the baseball video, wherein a start of at least one of the segments is first identified and then an end of the segment is identified based upon detecting a plurality of subsequent scene changes, wherein a first scene change detected after the start of the segment is free from being selected as the end of the segment, wherein the scene

changes are free from using a model sequence of frames, and wherein a subsequent scene change detected after the first scene change is selected as the end of the segment.

Claims 34-35 depend from claim 33 and are, therefore, patentable for at least the same reasons asserted for claim 33.

**Claims 36-38, 39 40-41, 44-45**

Claims 36-38 and 40-41 stand rejected under 35 USC 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (“Kawashima”) in view of Wang et al. (“Wang”) and further in view of Dimitrova et al. (“Dimitrova”).

Claim 36 patentably distinguishes over Kawashima in view of Wang and further in view of Rui by claiming identifying a plurality of segments of the video based upon said identifying as a result of creating a modified continuous segment by including a portion of the video identified during the sufficiently short duration, an identified segment immediately proceeding the portion, and an identified segment immediately subsequent to the portion, where the sufficiently short duration is greater than one frame.

There is no suggestion in these references to create a new segment as including the region not identified as a play. Rather it is presumed that the segment identified by Kawashima is in fact the entire play.

Claims 37-38 depend from claim 36 and are, therefore, patentable for at least the same reasons asserted for claim 36.

Claims 40-41 depend from claim 39 and are, therefore, patentable for at least the same reasons asserted for claim 39.

**Claims 39, 44 and 45**

Claims 39, 44, and 45 stand rejected under 35 USC 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (“Kawashima”) in view of Dimitrova et al. (“Dimitrova”).

Claim 39 patentably distinguishes over Kawashima in view of Dimitrova by claiming identifying a plurality of segments of the video, identifying a plurality of segments of the video, where each of the segments is identified as including a play of baseball, such that a plurality of segments are identified by the system as including a play of baseball; identifying at least one of the segments identified as including a play of baseball that has a temporally sufficiently short duration; wherein the temporally sufficient duration is greater than one frame, based upon the identifying removing the identified segment from the plurality of segments; and creating a summarization of the video by including said plurality of segments without the removed segment.

Claims 40-41 depend from claim 39 and are, therefore, patentable for at least the same reasons asserted for claim 39.

Claim 44 patentably distinguishes over Kawashima in view of Wang by claiming identifying a plurality of segments of the video wherein each of the segments includes a play of baseball, identifying a commercial within the video, and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video, where at least one of the plurality of segments of the video that is determined to be sufficiently temporally close to the commercial is not included within the summarization but would have otherwise been included within the summarization.

There is no suggestion in the references to remove an otherwise valid segment based upon detecting a commercial.

Claim 45 depends from claim 44 and is, therefore, patentable at least for the same reasons asserted for claim 44.

### **Claims 42 and 43**

The Examiner rejected claims 42 and 43 as being unpatentable over Kawashima et al. in view of Wee et al.

Kawashima discloses a method for summarizing a video whereby the start of a segment is of an at bat sequence (page 872, section 1.2). The process to find the at bat sequence necessarily involves model-based matching and action spotting (page 872, Figure 1). However, the model-based approach of Kawashima does not disclose anything whatsoever related to processing that

characterizes the relative position of a batter and whether the batter is sufficiently close to at least one of a catcher and an umpire proximate home base as claimed in claim 42.

The Examiner suggests that Wee is related to processing of regions being relatively close to one another. However, there is no mechanism within a model based system of Kawashima to incorporate such a relativity measure, nor any suggestion that one would be of benefit.

Claim 42 patentably distinguishes over Kawashima by claiming processing that characterizes a relative position of a batter and whether the batter is sufficiently close to at least one of a catcher and an umpire proximate home base. Kawashima fails to include such a characterization.

Claim 43 depends from claim 42 and is, therefore, patentable for at least the same reasons asserted for claim 42.

#### **Claim 46**

Claim 46 stands rejected under 35 USC 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (“Kawashima”) in view of Pan et al., and further in view of Lin et al.

Claim 46 patentably distinguishes over the referenced art by claiming, in part, identifying a plurality of segments of the video wherein each of the segments includes a play of baseball, wherein the identifying is in a manner free from using a model sequence of frames to determine the approximate frame that the play starts, wherein the segments include full-speed plays and slow motion plays of the full-speed plays, and further wherein the identifying is in a manner free from using an action spotting technique and free from using a pixel-by-pixel comparison.

The Applicants respectfully submit that all the pending claims should now be in a form suitable for allowance, and allowance of such claims is requested. If the Examiner for any reason believes that direct communication would advance the prosecution of this case, the Examiner is invited to contact the Applicants’ counsel at the number below.

Appl. No. 09/934,004  
Amdt. dated Apr. 25, 2006  
Reply to Office Action of Dec. 15, 2005



This Amendment is being submitted with a Petition for Extension of Time, together with the requisite fee. The Commissioner is hereby authorized to charge any additional fees, or credit any overpayment, to Deposit Account No. 03-1550.

Respectfully submitted,

CHERNOFF, VILHAUER, McCLUNG & STENZEL

Dated: April 25, 2006

By

Kevin L. Russell  
Reg. No. 38,292  
1600 ODS Tower  
601 SW Second Avenue  
Portland, OR 97204  
Tel: (503) 227-5631

**CERTIFICATE OF MAILING**

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on April 25, 2006.

Dated: April 25, 2006

Kevin L. Russell